

The Search for Truth: In Appreciation of James J. Heckman^{*}

July 2001

John J. Donohue
Stanford Law School

Stanford Law School
John M. Olin Program in Law and Economics
Working Paper No. 220

This paper can be downloaded without charge from the
Social Science Research Network Electronic Paper Collection:
http://papers.ssrn.com/paper.taf?abstract_id=276996

^{*} I am very grateful to Peter Siegelman for his excellent comments on an earlier draft of this paper.

The Search for Truth: In Appreciation of James J. Heckman
John J. Donohue*
July 2001

“On behalf of all economists who analyze micro-economic data and use micro-econometrics to unite evidence and theory, and to evaluate policy interventions of all kinds, I accept my portion of the Bank of Sweden Prize in Economic Sciences in memory of Alfred Nobel for the year 2000.”

--James J. Heckman, Dec. 8, 2000¹

The single most important advance in the social sciences in the last 25 years has been the enormous improvements in the ability to analyze microdata in order to identify and quantify causal relationships. Some of these gains have come from technological progress, as the vast increases in computing power have enabled the analysis of larger data sets using more sophisticated statistical techniques. Some of these gains have been organizational, as government and private entities have funded the collection of an extensive array of data sets that provide fertile grounds for academic researchers. But the most intellectually intriguing developments have been scientific, as the tools and methods of statistical analysis have been developed and sharpened. The immensely broad range and value of the applications of these tools across all empirical branches of social science, as well as in medicine, mandated recognition of this momentous scientific advance from the Royal Swedish Academy of Sciences. The choice of James Heckman as one of the two recipients of the 2000 Nobel Prize in Economics honors one of the giants in forging these new econometric techniques and modes of analysis.

I. The Need for and Difficulty in Predicting the Counter-Factual Post-Intervention World

The opening sentence of Jim’s Nobel lecture, quoted above, underscores a major goal of his scholarly work: to evaluate policy interventions by analyzing data in a way that unites evidence and theory. The difficulty of the task is frequently unappreciated by scholars, the public, and by policy makers, who often have the naïve belief that a “study” will clearly reveal the impact of any legal or policy change. But the daunting nature of uncovering complex causal relationships from non-experimental data is captured by the great Chicago economist Frank Knight, whom Heckman quoted in his Nobel Lecture:

“The existence of a problem of knowledge depends on the future being different from the past while the possibility of a solution of the problem depends on the future being like the past.”² A simple example will illustrate the nature of the difficulty. Figure 1 shows the time path of crime before and after the adoption of a hypothetical anti-crime measure in year 1985. Prior to 1985, crime had been rising but at a decreasing rate, and after year 1985 crime has trended down. In all likelihood, someone -- whether a policy maker who advocated adoption of the anti-crime measure or an ideological researcher who finds the policy to be compatible with his or her world view -- will want to argue that the anti-crime measure is responsible for the post-1985 drop in crime (and for similar reasons others will want to argue that it didn’t). Hopefully, there will also be a Heckmanesque researcher who doesn’t want to argue any position but simply wants to know whether the policy measure had any impact on crime.

A number of options are available to this truth-seeking researcher: for example, she can assume that absent the new policy, crime would continue to move as it did in the recent past by fitting a quadratic curve on the pre-1985 data. In this event, if the future is assumed to be like the past, the extension of the fitted quadratic would lead to the prediction that crime would have started falling some time after year 1985 in the absence

* I am very grateful to Peter Siegelman for his excellent comments on an earlier draft of this paper.

¹ This was the opening statement of Jim’s Nobel lecture. James J. Heckman, *Microdata, Heterogeneity and the Evaluation of Public Policy*, lecture presented at the Bank of Sweden Nobel Memorial Lecture in Economic Sciences, Stockholm, Sweden, December 8, 2000; webcast available from <http://www.nobel.se/economics/laureates/2000/heckman-lecture.ram>.

² Knight, Frank H., *Risk, Uncertainty and Profit* (Boston, New York: Houghton Mifflin Company, 1921).

of the policy intervention. Accordingly, this approach reduces the likelihood of a finding that the anti-crime measure mattered: crime would have been falling in any event, and only an unusually sharp decline (of greater magnitude than that shown in Figure 1, which represents a continuation of the pre-1985 quadratic trend) would be likely to generate a statistically significant finding that the policy reduced crime.

Alternatively, the researcher might fit a linear curve for the pre-adoption period. Now an extrapolation based on this linear curve will implicitly assume that, without the adoption of the policy in year 1985, crime would have continued rising, as shown in the post-1985 straight line counter-factual forecast in Figure 2. With this approach, the conclusion that the policy measure reduced crime seems to be powerfully buttressed by the evidence. But, as the quote from Knight suggests, and Heckman underscores without knowing whether, in the absence of the policy change, the future would be like the past, we can't know the effect of the legal change.

II. Becoming Attentive to the Issues of Selection Bias and Unobserved Heterogeneity

In some ways, the methodological advances that Heckman and others have pioneered over this last quarter century have been so successful that it is now difficult even to imagine a time when researchers were not attentive to concerns over selection bias and unobserved heterogeneity. Throughout all rigorously empirical branches of inquiry, these concepts have become familiar and frequently used – at least by the more sophisticated researchers in the various fields. The good news, then, is that the scientific contributions that the Nobel committee honored have substantially raised the quality of empirical research. State of the art academics have become far more sophisticated about the nature of observational (as opposed to randomized experimental) data, and alert to the dangers lurking in incomplete or selected data and the perils of spurious associations that masquerade as causal links.

Take the simple case of the simple-minded researcher who is interested in knowing whether Catholic schools are particularly successful in educating inner city blacks, as some have contended. To test this hypothesis, the researcher collects data on the SAT scores for inner city blacks, and runs a regression that includes a dummy to identify whether the student attended a Catholic high school or a public school. The resulting positive and statistically significant coefficient on the dummy for “Catholic school” is then interpreted as showing that blacks do better in Catholic schools than they do in public schools. Emboldened by her findings, the researcher announces that her analysis provides strong support for the view that vouchers enabling students to shift from public to private schools will lead to improved academic performance for blacks. Of course, the cognoscenti will now immediately perceive that the regression cannot provide an adequate basis for this conclusion. Attending Catholic school is not a condition randomly assigned across black inner city students, but is rather a choice variable. It is likely that the attributes of the black inner-city parents that chose to send their children to Catholic schools differ systematically from those of the average inner-city black parent. Indeed, one ordinarily suspects that low-income parents willing to pay for private school tuition might well value education more highly and provide an environment for their children that provides greater than average support for high academic achievement. Even if one could control for parental education and other measures of SES, traits unobservable by the researcher may vary systematically in a way that would tend to boost the SAT scores of the black children in Catholic schools, even apart from any added benefit over public schools that such private schools provide. This simple example illustrates the concepts of selection effects and unobserved heterogeneity, and in clarifying and highlighting the dangers posed by these effects and in providing sophisticated approaches to correct for them, Heckman has been a pioneer in enhancing the quality of estimates of treatment effects and policy interventions.

This is indeed very good news. The sophisticated producers of the various types of policy or treatment analyses will be fully cognizant of such issues and will strive mightily to discern truly causal links by using an array of statistical techniques or by fashioning an experimental design or identifying “natural” experiments. While the victory of these ideas among knowledgeable researchers has been complete, a broad canvass of current empirical research can be somewhat disheartening.

III. More Pearls or More Swine?

There are two ways in which the overall picture of microeconomic analysis *in practice* looks rather bleak. First, while one would like to report that studies such as the one caricatured above in which the researcher used a simple OLS regression to conclude that vouchers would generate academic gains for students have been eliminated from the landscape, the sad truth is that we are virtually drowning in studies of this kind.³ The technological and organizational factors alluded to above that have made high powered computing and large data sets widely available, coupled with the institutional efforts of the unscrupulous or untutored to offer empirical support for various policy measures, has enormously expanded the number of empirical studies of dubious quality. While the number of studies of the first rank has increased dramatically, the increase in the far larger number of virtually worthless studies conducted by the unknowledgeable or the devious has been a major setback. Going from 10 great empirical studies a year to 200 constitutes great progress, but going from 100 worthless studies a year to 1,000 breeds an often well-deserved cynicism about the value of empirical research, even though the percentage of valuable studies has risen considerably. In addition to the innovative techniques that the Nobel committee recognized in its award, one of the major contributions that Heckman has made to science has been as a model of excellence in crafting empirical research. It is to be hoped that the greater prominence his work will enjoy owing to his receipt of the Nobel prize will serve as a form of penicillin attacking the thriving bacteria of shoddy empirical work.

Second, there is an intermediate category of empirical studies that lies between the truly valuable and the utterly worthless. In this category falls studies done by those who have been trained in the more sophisticated techniques that Heckman and others have pioneered, and who have been lulled by their adherence to a set of statistical tools into thinking that they have reliably uncovered true causal relationships, when in fact they have largely uncovered numerical noise. The researchers in this class may have learned the first lesson about the significance of selection effects and unobserved heterogeneity and other land mines imperiling good statistical work, and they may have even learned the second lesson of following the cook book instructions of how to correct for these problems with advanced statistical techniques. But these more sophisticated techniques are not without cost. Their increased complexity means that it will be harder for researchers and readers to have a feel for the data, and it will be more difficult to critique an ostensibly well-crafted empirical study that isn't marred by one of the obvious pitfalls. As econometrics becomes increasingly technical, the rewards from the academy will more often go to those who have mastered complex mathematical techniques, rather than those who are alert to the relatively low-grade (but often pivotal) issues of data quality. Studies now pour out with conclusions based on increasingly fancy statistics, which their authors accept uncritically and which reviewers are at a loss to appraise effectively absent an enormous amount of work through attempted replication. The result at times mimics the title chosen by the famed Austrian economist Friedrich von Hayek for his 1974 Nobel lecture: *The Pretence of Knowledge*.

Although the technical sophistication of many recent empirical studies reveals that researchers have learned a great deal, the scholars of the second rank have not grasped one of the most important lessons embodied in Heckman's overall body of work: the truth is incredibly difficult to uncover and it cannot be assured by simply following a state of the art statistical protocol. This statement should certainly not be confused with the post-modern claim that there is no truth. The truth exists in the way that gold may exist in the ground even if miners have not yet discovered it, but the truth will never be found easily by following a pat set of

³ For example, during the Reagan Administration, a regression analysis purporting to show that SAT scores were actually *lower* in the states that spent the most on educating high school students was offered as evidence that cutting education budgets would not impair academic performance. The study failed to note that the primary reason that the SAT score was low in a high-spending state like Connecticut but high in a low-spending state like North Dakota was that dramatically more high school students took the test in the former state than in the latter. In effect, the average student in Connecticut was being compared with the very best student in North Dakota, thereby undermining the comparison that the researchers were trying to make (assuming that their goal was to uncover the truth as opposed to simply provide a smoke screen study to support a preferred policy option).

instructions based on instrumental variables, or corrections for selection effects, or extreme bounds analysis. These are tools, just as a computer is a tool, but, like the computer, they cannot simply be turned on with the hope that the truth will then be revealed. Jim Heckman has seen this more clearly and more constructively than virtually anyone who has ever lived. His work stands as a testament that the indispensable elements of any successful empirical research agenda include not only knowledge of the precise statistical techniques that Heckman and others have pioneered, but some less concretely identifiable skills of creativity, sound and tutored judgment, relentless effort to find and create more and better data, and exacting care in all phases of execution. By “successful” I mean a body of research that in fact advances our understanding of important social and economic truths, as opposed to success in the more crass personal sense or success in advancing one’s pre-existing ideology (the first of which is often orthogonal to my definition of a successful research program and the second of which is often directly opposed to it).

Heckman’s manner of letting truth trump ideology makes him the ideal scholar, and sets him apart from other University of Chicago Nobel luminaries such as Milton Friedman and George Stigler, who, despite their very important contributions to economics, were at times ruled by their rigid conservative ideologies. Moreover, the message was clear in the Chicago economics department of the 1970s and 1980s that studies conflicting with the conservative orthodoxy would not be warmly received. Under the watchful and approving eyes of Friedman and Stigler, a vast sea of Chicago school studies were issued by graduate students and assistant professors showing the ineffectiveness or harmfulness of governmental interventions.⁴ Against this monolithic negative assessment of government stands the work of one young Chicago school professor who was sufficiently committed to the truth, and intellectually powerful and independent enough to resist the prevailing orthodoxy to generate an extraordinary body of work on the impact of federal civil rights laws on the employment prospects of southern blacks. While the official announcement of the Nobel prize stated that “James Heckman is the world’s foremost researcher on econometric policy evaluation,”⁵ it should be remembered that it took a scholar of unyielding power and integrity to defy the views of the then Chicago giants. It is perhaps for this reason that Jim told the Wall Street Journal upon winning his Nobel prize that the work of which he was most proud was his evaluation of the impact of Title VII.⁶

IV. Heckman’s Study of the Impact of Title VII on Black Employment in the Southern Textile Industry

In an article entitled “Determining the Impact of Federal Antidiscrimination Policy on the Economic Status of Blacks,” (*The American Economic Review*, Vol. 79, No. 1 (Mar., 1989), pp. 138-177), Heckman and coauthor Brook Payner show that between 1910 and 1964, black employment was low and stable in the textile industry⁷ in South Carolina; in 1965, however, there was a sudden jump in the employment rate and wage levels of both male and female black workers in this industry. This change, Heckman and Payner argue, is attributable to the implementation of Title VII of the 1964 Civil Rights

⁴ Many of these articles have not stood the test of time: for example, despite the interesting theoretical insight found in Sam Peltzman’s famous paper suggesting that auto safety legislation was responsible for increased pedestrian deaths and more overall accidents, the dramatic drops in both per mile motor vehicle occupant and pedestrian deaths over the last thirty years would seem to belie the empirical validity of Peltzman’s work. See Henry Manne and Roger Miller, eds., *Auto Safety Regulation: The Cure or the Problem?* 65-66 (1976) for Peltzman’s thesis and the devastating critique by Richard Nelson.

⁵ P5

⁶ Steve Liesman, “Two U.S. Economists Win Nobel Prize,” *The Wall Street Journal* p. A2 (October, 12, 2000).

⁷ In this article, Heckman and Payner’s broad focus is on the traditional manufacturing sector of South Carolina, which was already well established when Jim Crow laws formalized racial segregation in the early 1900s. Within this sector, Heckman and Payner focus more specifically on the textile industry because it is a major industry in the state that accounted for the majority of all manufacturing employment in 1940 and a sizable percent in 1980. In addition, Heckman and Payner write, “Title VII and related antidiscrimination activity seems likely to have had its most visible effect in industries like South Carolina textiles in which exclusion of blacks was so blatant.” (p. 143).

Act and other related federal antidiscrimination policies that went into effect in 1965. In reaching this conclusion, Heckman and Payner carefully assess and reject several major counter-explanations for this economic progress of blacks. What emerges is an article that is a paragon of careful and meticulous scholarship that not only provides thorough empirical support for its main thesis, but displays keen sensitivity to various counter-explanations. Below is a brief synopsis of some of the points that Heckman and Payner make in their article, which illuminate their nuanced and comprehensive approach to analyzing this issue.⁸

Heckman and Payner first present five striking graphs of wage and employment trends in South Carolina by race and sex between 1910 and 1980, which are based on the data from the Annual Reports of the South Carolina Department of Labor. As the article notes:

Through two World Wars, the Great Depression, and the booms of the 1950s and 1920s the share of blacks in textile employment remains constant at a low level despite growth in the quality and quantity of black schooling and despite economic scarcity resulting from tight labor markets. Suddenly in 1965 the black share in employment begins to improve when Title VII legislation becomes effective and the Equal Opportunity Commission begins to press textile firms to employ blacks and when Executive Order 11246 forbids discrimination by government contractors at the risk of forfeiture of government business. Textiles sold a significant proportion of their output to the federal government in 1965. The improvement in black employment and wages occurs at a time and in an industry that suggests a major role for governmental activity. (p. 143.)

Fifty-five years of data showing virtually no blacks in the South Carolina textile industry followed by very substantial increases in 1965 created powerful visual evidence that the federal government antidiscrimination policy improved the economic status of blacks starting in 1965. However, rather than simply depend on these graphs to tell the story, Heckman and Payner carefully analyze various potential arguments against their conclusion. The first criticism they anticipate is the notion that their underlying data might be inaccurate because the employers might have had incentives to lie about the black employment rate after 1965 to avoid federal intervention. To counter this argument, Heckman and Payner carefully compare the South Carolina Department of Labor data with the U.S. Census of Population data, which yield the same pattern of dramatic black improvement beginning in the year 1965.

Heckman and Payner address several other salient potential counter-arguments to their government activity thesis – specifically, two supply-side arguments and one demand-side argument. One such supply-side argument that Heckman and Payner discuss is based on the theory that the decline in agricultural labor during the 1950s led to shifts in the supply of blacks to the manufacturing sector. Heckman and Payner, however, point out that this theory is inadequate to explain the black breakthrough in the textile sector because a study of employment trends by sex show that most of the increase in black manufacturing employment came from black females, whereas their agricultural employment saw only a relatively small decline. Moreover, Heckman and Payner point out that in general “only a tiny fraction of black entrants into industry came from agriculture.” (p. 149).

The second supply-side argument that Heckman and Payner analyze and reject is the argument that an increase in the quantity and quality of black education during the 1940s and 1950s allowed blacks to better compete with whites in the labor market. Using a series of linear probability regression models of black educational attainments and employment trends, however, Heckman and Payner succeed in showing that “controlling for education and other individual characteristics, blacks are significantly underrepresented in the textile industry in 1960 and before. In 1970, however, the underrepresentation disappears for the younger cohorts and diminishes considerably for the older cohorts. If educational improvement led to the black gains in textile employment, the regressions controlling for education would have shown no underrepresentation of blacks in 1960, and no change in underrepresentation from 1960 to 1970.” (p. 158).

⁸ Heckman and Payner’s government activity hypothesis takes on an added significance, as they explain that trends in black economic progress in South Carolina typify trends for the South as a whole, which in turn, account for the national trend in black economic progress during the period between 1960 and 1980.

They also carefully point out that textiles is a low skill industry, and that increases in black education should have reduced, not increased, black employment in this industry between 1960 and 1970.

Finally, Heckman and Payner turn their attention to the tight labor market hypothesis, according to which a strong economy and low unemployment rate in the mid to late 1960s resulted in an increased pressure for employers to integrate the workforce as means of keeping down labor costs. Faced with this counter-explanation, Heckman and Payner conduct a careful examination of eleven different South Carolina counties to determine whether effects of changing market conditions have varied across different counties. This analysis is based on the assumption that since South Carolina counties do not form a single labor market, then a tight labor market would affect the counties differently, whereas uniformly applied federal government policy of antidiscrimination would have a simultaneously parallel effect on all the counties. Their study of the eleven counties, indeed, show simultaneous, rather than varied, breakthroughs of black employment in the mid-1960s for each of the counties.⁹ However, even in face of this data, Heckman and Payner adopt a more nuanced approach by leaving open the possibility that “in 1965 entrepreneurs seized on the new federal legislation and decrees to do what they wanted to do anyway” or that perhaps “the federal antidiscrimination and affirmative action laws came into existence in 1964 precisely because the U.S. labor market was tight to an unprecedented degree and discrimination was becoming costly.” One of the great features of this study is that it doesn’t stop when the data seems to confirm the author’s preferred theory, but goes to great lengths fairly to consider and rebut all of the most-plausible counter arguments. While I don’t quite understand all of the dimensions of the dispassionate, powerful, and disciplined mind and will needed to produce this type of work, the remarkable infrequency of this approach testifies powerfully that these traits are extremely rare.

V. Conclusion

Those of us who have learned from Jim Heckman by watching and assisting in his relentless search for the truth were delighted to see his pathbreaking work acknowledged to the world at large through his receipt of the Nobel prize in the year 2000. He has been the exemplar of the researcher constantly vigilant against the bias of ideology, which so often leads scholars to uncritically accept studies whose conclusions they like while painstakingly critiquing those whose conclusions they dislike. Jim challenges everything with the tutored skepticism that is the hallmark of the scientific search for truth. In an age when partisanship and ideology, perhaps encouraged by the strange winds of post-modernism, have risen to such an unhealthy level that they taint virtually all policy discussions in the political world and vast realms of academic writing, it is a godsend that the Nobel committee has honored one of the true giants in the search for truth. Jim Heckman is in the rarefied world of the most knowledgeable handful of people who have ever lived in terms of his understanding of econometrics and statistics. He knows more history and math than any other economist. And again is among a small handful of the greatest empirical economists and quantitative social scientists that has ever lived. Combine these all into one person, and you get quite an intellectual force. Heckman has shown how difficult it is to know the truth without rejecting knowing the truth as a goal. What Beethoven was to music, Heckman has been to economics – the wild, creative genius who having arrived on the scene leaves the world not quite the same ever again.

On a personal note, I confess that, just as spent years watching and trying to mimic, admittedly with imperfect success, the basketball game of the great Knick guard Walt Frazier, I have shaped my approach to research on that of Jim Heckman. He is truly one of the giants in all of social science, and I feel that I have profited enormously from the opportunity to work and learn from this towering figure of the academy. Jim has always been the rebel – brilliant and uncompromising – who has pursued the truth with an intensity that few have ever matched. While those who know him appreciate how astoundingly funny he can be, I know I am not revealing any secret when I acknowledge that Jim is not always the easiest person to deal with. But his intense demands for the near impossible only reflect the intense demands he places on himself and must be understood as emanating from his passion for the truth, his relentless determination to

⁹ The study examines the eleven counties “for which we can form a consistent time-series on textile employment by race and sex for the period 1910-77.(p.160). They also explain that the counties vary greatly in racial composition, size of the population, size of manufacturing work force, and thus “[t]o the extent that regional labor markets exist, we expect the effect of changing labor market conditions to be different across counties.”

uncover it, and the vast scope of his intellectual horizons. With appreciation and admiration, I encourage all scholars to familiarize themselves with Jim Heckman's impressive body of scholarship and learn from his impeccable scholarly integrity.

Figure 1:
Crime Rates Over Time
Actual Hypothetical Crime Rates and Fitted Quadratic Equation

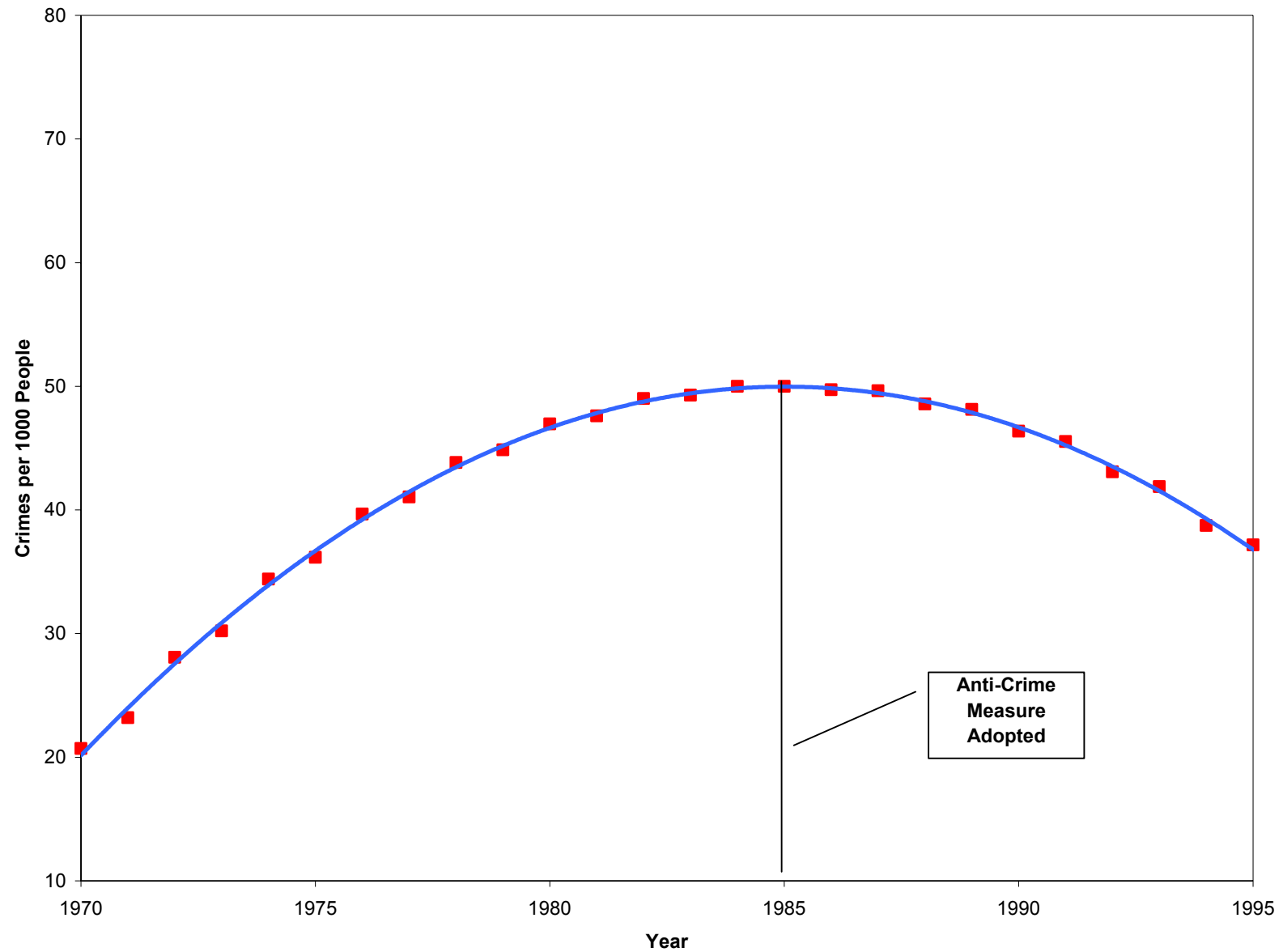


Figure 2:
Crime Rates Over Time
Actual Hypothetical Crime Rates and Linear Pre-Passage Trend Line

